

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
15 January 2004 (15.01.2004)

PCT

(10) International Publication Number
WO 2004/005933 A1

(51) International Patent Classification⁷: **G01N 33/68**,
C07K 1/13, C09B 23/02, 67/22

[GB/GB]; Amersham Biosciences UK Ltd, The Grove
Centre, White Lion Road, Amersham, Buckinghamshire
HP7 9LL (GB).

(21) International Application Number:
PCT/GB2002/003142

(74) Agents: **FRANKS, Barry** et al.; Amersham plc, The
Grove Centre, White Lion Road, Amersham, Bucking-
hamshire HP7 9LL (GB).

(22) International Filing Date: 8 July 2002 (08.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (for all designated States except US): **AMER-
SHAM BIOSCIENCES UK LIMITED** [GB/GB]; Amer-
sham Place, Little Chalfont, Buckinghamshire HP7 9LL
(GB).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WILLIAMS, Karen**
[GB/GB]; Amersham Biosciences UK Limited, The Grove
Centre, White Lion Road, Amersham, Buckinghamshire
HP7 9LL (GB). **STONE, Timothy** [GB/GB]; Amersham
Biosciences UK Ltd, The Grove Centre, White Lion
Road, Amersham, Buckinghamshire HP7 9LL (GB). **SIM-
MONDS, Adrian, Christopher** [GB/GB]; Amersham
Biosciences UK Ltd, The Grove Centre, White Lion Road,
Amersham, Buckinghamshire HP7 9LL (GB). **SWEET,
Alison, Claire** [GB/GB]; Amersham Biosciences UK Ltd,
The Grove Centre, White Lion Road, Amersham, Buck-
inghamshire HP7 9LL (GB). **FOWLER, Susan, Janet**

(54) Title: REAGENTS AND A METHOD FOR SATURATION LABELLING OF PROTEINS

(57) Abstract: A matched set of fluorescent dyes is provided, wherein each dye of the set is capable of covalent attachment to a protein and wherein each of the dyes has a molecular structure and a charge that is matched one with the other, such that relative electrophoretic mobility of a protein labelled with one dye of the set is the same as the electrophoretic mobility of the protein labelled with a different dye of the set. The matched set comprises at least two different fluorescent dyes of formula: wherein n is 1, 2, or 3; Z1 and Z2 independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system; one of groups R1 and R2 is a target bonding group; remaining group R1 or R2 is selected from -(CH₂)₄-W or -(CH₂)_r-H; group R3 is hydrogen, except when either R1 or R2 is -(CH₂)_r-H, in which case R3 is W; and W is selected from sulphonic acid and sulphonate. The invention also provides a method for saturation labelling of a protein with a fluorescent dye so as to label all available target amino acid, suitably cysteine, residues in the protein, thereby giving a single population of labelled protein molecules.



WO 2004/005933 A1

Reagents and a Method for Saturation Labelling of Proteins

The present invention relates to fluorescent dyes and to a method for labelling complex protein samples to enable differential protein
5 analysis.

2D Difference Gel Electrophoresis (DIGE) uses matched, spectrally-resolved fluorescent dyes to label protein samples prior to 2-dimensional (2D) electrophoretic separation (Minden, J. et al, Electrophoresis, (1997),
10 18, 2071). This fluorescent multiplexing approach overcomes many of the disadvantages of traditional 2D electrophoresis. Fluorescent pre-labelling of protein samples allows multiple samples to be run on the same gel, enabling quantitative differences between the samples to be easily identified by overlaying the fluorescent images. To allow fluorescent
15 multiplexing, the protein samples must be labelled equally with the dyes and the migration of the labelled proteins in the gel must be positionally matched, to allow quantitative differences to be detected.

WO 96/33406 (Minden, J. et al) discloses a method to detect
20 differences between different cell samples using matched, spectrally resolved dyes to label protein in the samples prior to 2D electrophoresis. Described are dyes that are matched for molecular mass and charge to give equivalent migration. The approach employs cyanine dyes having an N-hydroxysuccinimidyl (NHS) ester reactive group to label amines. The
25 NHS ester group is targeted at the ϵ -amino of lysine residues in proteins and results in covalent labelling. Three different dyes (Cy2TM, Cy3 and Cy5) are derivatised to enable multiplexing. These dye molecules have a molecular weight of approximately 500 Daltons and are matched in mass to give equivalent migration of the labelled protein. The dyes rely on the
30 intrinsic charge of the cyanine fluor to compensate for the loss of a lysine positive charge on conjugation of the dye to lysine. WO 96/33406

describes the use of these dyes in a minimal labelling strategy to label between 1 and 2% of available attachment sites. To achieve equivalent migration, the dye pairs were matched for mass by compensating for the difference in the linker length between the indole rings of the cyanine by modulation of the size of the aliphatic chain attached to an indole nitrogen by 2 carbon units. Two or more spectrally resolved dyes that meet these criteria represent a matched minimal labelling dye set. Equivalent migration of the matched lysine minimal labelling dye set was demonstrated using propyl Cy3 and methyl Cy5 to label proteins followed by their 2D separation and overlaying the resulting images to show positional matching of the two labelled samples.

Lysine residues are highly abundant in proteins, ensuring that this labelling strategy will represent all the proteins present in a complex sample. However, the typical lysine content of a protein is 7% and, if every lysine in a protein were labelled, it would result in a large mass shift due to the dye. Thus, the strategy employed is to "minimally" label the protein to ensure that only ~1 in 5 protein molecules are labelled, thereby giving a statistical probability that each labelled molecule has only one dye attached. This creates a spot pattern on 2D that is very similar to silver stained images, but gives greater sensitivity and dynamic range and the ability to multiplex.

A typical 2D gel analysis involves "picking" protein spots of interest from the gel for identification by MALDI-MS. The minimal labelling approach results in 2 spots per protein (ie. labelled and unlabelled), with the majority of protein in the unlabelled spot. The unlabelled spot must be located to recover sufficient protein to enable identification by MALDI-MS. This requires an additional staining step prior to spot picking. In order to facilitate protein spot picking directly from fluorescent gels, a labelling strategy is required that gives a single spot per protein.

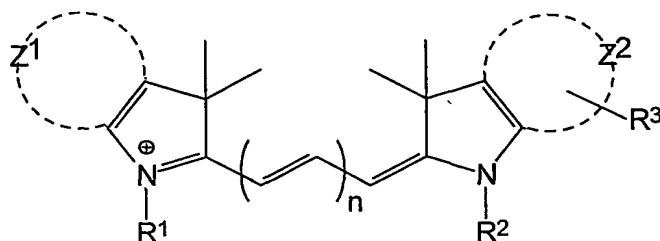
The labelling strategy employed in the present invention aims to saturate the protein with dye in order to label as many target residues as possible and to produce a single labelled spot per protein isoform. Thus, a cyanine dye molecule is coupled to all available target amino acid residues on a protein, thereby giving a single population of labelled protein molecules with a similar number of dye molecules attached. According to the present invention, saturation labelling is achieved using dye sets targeted at cysteine residues, which contain a thiol group. Cysteine residues are present in 95% of proteins, but there are fewer cysteine residues in each protein than lysine. This means that all protein molecules can be labelled, but each protein molecule will have fewer labelled residues than if lysine residues were labelled to saturation. This results in increased sensitivity of detection as the proportion of labelled protein is higher than with minimal labelling, but ensures that the proteins remain soluble.

The increase in mass due to the addition of dye (the mass shift) with saturation labelling of cysteine residues will be larger than that employing a minimal labelling strategy of lysine residues. However, the increase in mass is less than would be observed if a saturation labelling approach was employed on lysine residues. The extent of the mass shift due to the dye addition will vary for each protein depending on the individual protein cysteine content (typically ~2%) and the availability of the cysteine residues to the dye under the conditions of labelling. This results in a 2D spot pattern that is very different from published silver stained 2D protein maps.

The present invention therefore provides fluorescent reagents and a method for reproducibly labelling, all of the available cysteine residues that are accessible in a mixture of cysteine-containing proteins. The cyanine dye derivatives according to the present invention provide

valuable sets of fluorescent labels, each having a common core structure and which are particularly useful for multiplex analysis.

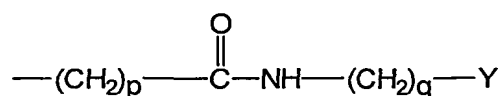
Accordingly, in a first aspect, the present invention provides a
 5 matched set of fluorescent dyes comprising at least two different
 fluorescent dyes having the formula (I):



(I)

wherein n is different for each said dye and is 1, 2, or 3;

15 Z¹ and Z² independently represent the carbon atoms necessary to
 complete a phenyl or naphthyl ring system;
 one of groups R¹ and R² is the group:



20 where Y is a target bonding group;

remaining group R¹ or R² is selected from $-(CH_2)_4-W$ or $-(CH_2)_r-H$;

group R³ is hydrogen, except when either R¹ or R² is $-(CH_2)_r-H$, in which
 case R³ is W;

25 W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

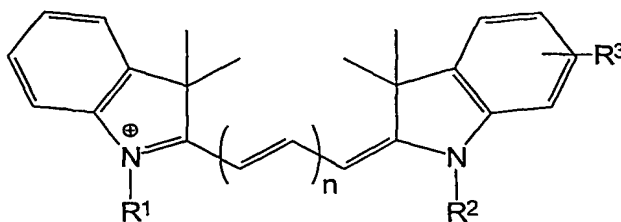
r is an integer from 1 to 5;

and their salts;

30 characterised in that when n of two of said dyes differs by + 1, one of p,
 q and r of said two dyes differs by -1.

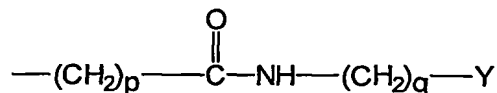
According to the present invention a matched set of fluorescent dyes is provided, wherein each dye of the set is capable of covalent attachment to a protein and wherein each of the dyes has a molecular structure and a charge that is matched one with the other, such that relative electrophoretic mobility of a protein labelled with one dye of the set is the same or substantially the same as the electrophoretic mobility of the protein labelled with a different dye of the same set. In one embodiment according to the first aspect, the matched set of dyes comprises at least two different fluorescent dyes, wherein each dye in said set is a compound having the formula (I) and wherein Z^1 , Z^2 , R^1 , R^2 , R^3 , Y , W , n , p , q and r are hereinbefore defined. Suitably, the different dyes in the matched set of fluorescent dyes are spectrally resolvable to enable different samples labelled with such dyes to be distinguished one from the other. Suitably, at least two dyes of the matched set of dyes have a structure according to formula (I) and may be selected from the trimethine cyanine dye class (in which $n = 1$), the pentamethine cyanine dye class (in which $n = 2$) and the heptamethine cyanine dyes (in which $n = 3$).

Suitably, the matched set of fluorescent dyes according to the invention comprises at least two different fluorescent dyes having the formula (II):



(II)

wherein n is different for each said dye and is 1, 2, or 3;
one of groups R^1 and R^2 is the group:



where Y is a target bonding group;

- 5 remaining group R¹ or R² is selected from $-(\text{CH}_2)_4-\text{W}$ or $-(\text{CH}_2)_r-\text{H}$;
 group R³ is hydrogen, except when either R¹ or R² is $-(\text{CH}_2)_r-\text{H}$, in which
 case R³ is W;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

- 10 q is selected to be 2 or 3; and

r is an integer from 1 to 5;

and their salts;

characterised in that when n of two of said dyes differs by + 1, one of p,
 q and r of said two dyes differs by -1.

15

Preferably, the matched set of fluorescent dyes comprises at least
 two different dyes according to formula (I) or (II) in which:

n is selected to be 1 or 2;

p is selected to be 4 or 5;

- 20 q is selected to be 2 or 3; and

r is selected to be 1, 2 or 3.

Preferably, the target bonding group Y in each dye of the matched
 set of fluorescent dyes is the same and is selected from a maleimido
 25 group and an iodoacetamido group. A particularly preferred target
 bonding group for each dye is a maleimido group.

Particularly preferred are dye sets according to formula (I) or (II)
 that are selected from the trimethine cyanine class of dyes and the
 30 pentamethine cyanine class of dyes. Such dyes are described as Cy3™

and Cy5 dyes. Absorbance and emission data for the cyanine dyes are shown below in Table 1.

Table 1

<u>Dye</u>	<u>Fluorescence Colour</u>	<u>Abs (nm)</u>	<u>Em (nm)</u>
Cy3	Orange	550	570
Cy5	Far red	649	670
Cy7	Near IR	747	774

Suitably, salts of the fluorescent dyes according to formula (I) or (II) may be selected from K^+ , Na^+ , NH_4^+ , R_3NH^+ and R_4N^+ where R is C_1 to C_4 alkyl.

In an alternative embodiment, the matched set of dyes may optionally include one or more additional dyes, provided that each such additional dye possesses charge and mass characteristics, such that the electrophoretic mobility of a protein labelled with the dye is the same or substantially the same as the electrophoretic mobility of a protein labelled with a dye according to formula (I) or (II). Other such additional dyes may include benzoxazole-containing dyes. An example of an additional dye is Cy2.

Exemplary sets containing pairs of fluorescent dyes according to formula (I) or (II) that are matched in electrophoretic mobility when coupled to proteins are as follows:

Set 1

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound I); and

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1,3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound II);

5 Set 2

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(1-propyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound III); and

10 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound IV);

Set 3

15

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound I); and

20 1-(5-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound V);

Set 4

25 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound VI); and

30 1-(5-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound VII).

Set 5

1-(6-{[3-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-
2-[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-
5 ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound VIII); and
1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-
3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-
2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound IV); and

Set 6

1-(6-{[3-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-
2-[(1*E*,3*E*)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-
ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound IX); and
15 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-
3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-
2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound X).

The present invention therefore provides a matched set of reagents
20 for reproducibly labelling all cysteine residues accessible in a mixture of
cysteine-containing proteins under the conditions used. Equivalent
migration in the pI dimension is achieved using dye sets with an overall
neutral charge in order to match the neutral charge on the thiol group to
which the dyes are conjugated. Neutral cyanine dyes may be obtained by
25 means of a sulphonic acid or sulphonate group substituent covalently
attached to the dye structure. Moreover, it has been discovered that
equivalent mass migration is not simply achieved by matching the
molecular weight of the dye sets, but rather, requires careful manipulation
of the overall size and mass of the dye in order to achieve positional
30 matching of proteins labelled with such dyes. In order to obtain
equivalent migration employing a cysteine saturation labelling approach,

dye sets are required wherein each of the matched dyes according to formula (I) differs in mass by a single carbon unit. The dye sets may be obtained by varying different substituents attached to the dye chromophore. Two or more spectrally resolved dyes that meet these
5 criteria represent a matched saturation dye set.

The present invention also provides a method for saturation labelling of a protein with a fluorescent dye so as to label all available target amino acid residues in the protein, thereby giving essentially a
10 single population of labelled protein molecules. Suitably, the target amino acid is a cysteine residue. By the term "available" it is meant amino acid residues that are accessible to the fluorescent dye for reaction. Available cysteine residues must be accessible, and in a reduced (i.e. in a free thiol) form.

15 Thus, in a second aspect, there is provided a method for labelling a mixture of proteins in a sample wherein each of said proteins contains one or more cysteine residues, said method comprising:

- i) adding to an aqueous liquid containing said sample a fluorescent
20 dye wherein said dye contains a target bonding group that is covalently reactive with said proteins; and
- ii) reacting said dye with said proteins so that said dye labels said proteins;

characterised in that all available cysteine residues in said proteins are
25 labelled with said dye.

Preferably, the fluorescent dye according to the method of the second aspect is a cyanine dye. Particularly preferred cyanine dyes for use in the method are those containing a sulphonic acid or sulphonate
30 group, for example, the dyes according to formula (I).

Preferably, the target bonding group is selected from a maleimido group and an iodoacetamido group.

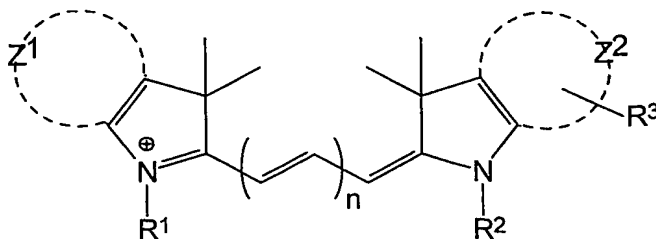
In a preferred embodiment according to the second aspect, the method further comprises prior to step i), the step of treating the protein with a reductant.

Preferably, the cyanine dye is used in a range of 5 to 200nmol of dye per 50 μ g of protein.

Preferably, the method for labelling a mixture of proteins in a sample according to the second aspect is performed at a pH in the range 6.0 to 9.0.

The invention also provides a method for labelling and thereby imparting fluorescent properties to a sample, suitably a protein sample, using a dye according to formula (I). Thus, in a third aspect, there is provided a method for labelling one or more proteins in a sample, the method comprising:

i) adding to a liquid sample containing said one or more proteins a fluorescent dye of formula (I):

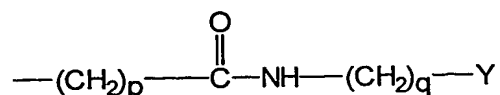


(I)

wherein n is different for each said dye and is 1, 2, or 3;

Z¹ and Z² independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system;
one of groups R¹ and R² is the group:

5



where Y is a target bonding group;

remaining group R¹ or R² is selected from $-(\text{CH}_2)_4-\text{W}$ or $-(\text{CH}_2)_r-\text{H}$;

10 group R³ is hydrogen, except when either R¹ or R² is $-(\text{CH}_2)_r-\text{H}$, in which case R³ is W;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

15 r is an integer from 1 to 5;

and their salts;

characterised in that when n of two of said dyes differs by + 1, one of p, q and r of said two dyes differs by -1; and

incubating said dye with said sample under conditions suitable for

20 labelling said one or more proteins.

Preferably, each of Z¹ and Z² represents the carbon atoms necessary to complete a phenyl ring system.

25 Preferably, the method according to the third aspect employs a fluorescent dye of formula (I) in which:

n is selected to be 1 or 2;

p is selected to be 4 or 5;

q is selected to be 2 or 3; and

30 r is selected to be 1, 2 or 3.

Preferably, the target bonding group Y is selected from a maleimido group and an iodoacetamido group.

To perform the method for saturation labelling of cysteine residues in proteins, a number of reaction parameters must be considered, including:

i) Accessibility of Protein Thiol Groups

In native folded proteins many cysteine residues are involved in the secondary structure of proteins by disulphide bond formation and may also be buried within the core of the molecule. In order to label these residues the protein must be unfolded to enable accessibility of the dyes, using protein denaturants such as urea, and they must be reduced to generate the free thiol. The choice of reductant is important to enable efficient reduction of the protein. Commonly used thiol-containing protein reductants are dithiothreitol (DTT) and β -mercaptoethanol. Suitable phosphine reductants include tributyl phosphine (TBP) and tris(2-carboxyethyl) phosphine (TCEP). Sodium borohydride may also be used. TCEP is a preferred reductant to give efficient reduction, whilst minimising reaction with the dye.

The efficiency of reduction is influenced by the concentration of reductant (relative to the cysteine content of protein sample), the temperature of reaction (which influences protein denaturation) and the duration of the reaction. If the reductant is immobilised on beads this may also affect the efficiency. The optimal reductant concentration for a particular sample type will vary depending on the reductant used and the cysteine content of the sample. However, 50 μ g of protein typically requires 10nmol of TCEP for reduction (assuming average cysteine content at 2%). For mammalian samples with higher cysteine content,

then 20nmol of TCEP is typically used. It is also important to maintain the ratio of 1nmol TCEP to 2nmol dye to give optimal labelling.

The optimum combination of these factors varies depending on the individual protein structure and cysteine content. Higher reduction temperature will give increased labelling, as a result of increased denaturation of the protein and thus increased accessibility of cysteines. However, high temperatures may have an adverse effect on the protein. Urea exists in equilibrium with carbamylate, but at higher temperatures (generally $>40^{\circ}\text{C}$) the equilibrium shifts in favour of the carbamylate. Carbamylate is a more chemically reactive species than urea and can attack primary amines (eg. N-terminal amino group and ϵ -amino group of lysines), leading to artefactual charge heterogeneity and the generation of charge trains on 2D gels. Saturation labelling of cysteine residues at 37°C in the presence of urea does not exhibit carbamylation effects.

ii) Dye Concentration

The hydrophobic nature of the cyanine dyes requires the use of organic solvents (eg. 10% DMF) to aid solubility. To achieve saturation of the available target residues so that the labelling goes to completion it is necessary to use high concentrations of dye for labelling, whilst maintaining dye solubility. The use of a sulphonate group on the cyanine dyes helps to maintain solubility at high dye concentrations. The optimal dye concentration in the labelling reaction for a particular sample type will vary depending on the cysteine content of the sample and whether there are any other components in the sample which might interfere with labelling. However, the dye should always be present at least in excess of the thiol content to achieve saturation labelling and it is also important to maintain the ratio of 1nmol TCEP to 2nmol dye to give optimal labelling. Using the dye in excess would typically require the dye to be in

the range 5 to 200nmol per 50µg of protein sample. However, 50µg of protein typically requires 20nmol of dye for a labelling reaction (assuming average cysteine content of 2%). For mammalian samples with higher cysteine content, 40nmol of dye is typically used. However, other components in the sample may also react with the dye. For example, liver samples may contain elevated levels of the tripeptide glutathione, in response to stress/toxic substances. As a consequence, increased dye concentrations may be required in order to maintain labelling efficiency. Serum samples containing elevated levels of albumin, which is a highly abundant protein with a high cysteine content, may also require increased dye concentrations.

iii) pH of Labelling Reaction

Cysteine residues are strong nucleophiles and may be labelled in proteins using reagents having iodoacetamide and maleimide reactive groups. A particularly preferred reactive group is the maleimido group. The pKa of a thiol group is critical in determining its reactivity. Above the corresponding pH, its nucleophilicity increases markedly as the thiolate (S⁻) form replaces the protonated SH species. In an isolated cysteine molecule, the thiol group has a pKa of approximately 8.6; however, the microenvironment of the thiol within the protein can also effect its pKa. The rate of reaction of thiol groups with maleimides increases at alkaline pH, due to the increased concentration of thiolate anion. Under alkaline conditions, hydrolysis of the maleimide to maleamic acid becomes a significant side reaction, and competing reactions with other functional groups such as lysine and histidine become more significant.

The ε-amino group of lysine behaves as a typical amine with a pKa in proteins of 9.0 – 9.5. At a lower pH, e.g. pH 6.5, the majority of amine groups are predominantly in the protonated, unreactive NH₃⁺ form

and the reaction with maleimide groups is ~1000 fold slower than with thiols. Terminal amine groups have pKa values of 7.5 – 8 depending on the amino acid concerned. The terminal amine will be present in the un-

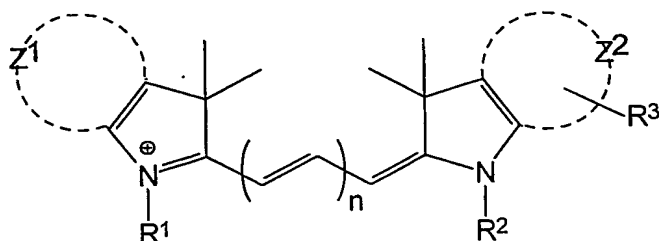
protonated (reactive) form at lower pH than those of lysine. Thus,
5 reaction of maleimide groups with amines requires a higher pH than the reaction with thiol groups. Consequently, labelling of thiol groups with maleimides is more specific at lower pH, with less potential to label lysine residues. Thus, the pH of the reaction is also a critical factor in achieving saturation labelling and for specific labelling of thiol groups. Preferably,
10 saturation labelling using fluorescent dyes, for example according to formula (I) is performed at a pH range of 6.0 to 9.0, most preferably a pH of 8.0, being a compromise between the presence of the reactive thiolate and the presence of reactive amines.

15 Protein samples for comparison may be derived from a variety of cell sources, including all normal and transformed cells derived from any recognised source with respect to species (eg. human, rodent, simian), tissue source (eg. brain, liver, lung, heart, kidney skin, muscle) and cell type (eg. epithelial, endothelial). There are established protocols available
20 for the culture of diverse cell types. (See for example, Freshney, R.I., Culture of Animal Cells: A Manual of Basic Technique, 2nd Edition, Alan R.Liss Inc. 1987). This invention may also be used to compare samples from plants using intact plants or cultured plant cells. In addition, samples for use in the method of the invention may be derived from cells
25 which have been transfected with recombinant genes and cultured, cells which have been subjected to an external stimulus (such as heat shock or drug treatment) or other biological fluids (such as cerebrospinal fluid or serum). The present invention may also be used to target subsets of proteins present in a cell prior to labelling; for example, by isolating
30 particular fractions such as low molecular weight proteins or pl range; sub-cellular compartments such as nuclear proteins.

Those skilled in the art will recognise that protein samples may be extracted from such samples using a variety of methods and extraction reagents. Typically, cells from the tissue/culture are disrupted, for example by homogenisation, sonication, cell lysis, and the protein extracted and solubilised in the presence of reagents including denaturing reagents, such as urea, thiourea, detergents such as SDS, CHAPS, Triton X-100, NP-40, reducing agents, such as dithiothreitol (DTT), mercaptoethanol, and buffer such as Tris, Hepes. Protease inhibitors, such as phenylmethanesulphonyl fluoride (PMSF), ethylenediaminetetraacetic acid (EDTA), leupeptin, aprotinin, may also be added to minimise degradation by endogenous proteases.

The matched set of fluorescent dyes and the labelling method may be used for detecting differences in the protein content of at least two different protein samples, for example, according to the method described by Minden, J. et al (Electrophoresis, (1997), 18, 2071). In a typical example of the method, protein is extracted from each of 2 different samples by known methods as described above and the protein concentration determined. To a 50µg aliquot of each protein extract is added 10nmol of TCEP (tris-(2-carboxyethyl)phosphine) and this is incubated for one hour at 37°C. To each of these reduced protein samples is added 20nmol of dye and incubated for 30 minutes at 37°C. The first protein sample is labelled with the first dye (for example Cy3) of a matched set of dyes and the second protein sample is labelled with the second dye of the matched set of dyes (for example Cy5). The reaction is stopped by the addition of sample buffer containing DTT. The labelled protein samples are mixed to give a single sample for separation. The samples are electrophoretically separated, preferably by 2D-PAGE. The procedures for electrophoretic separation are well known to those skilled in the art.

In a fourth aspect of the invention, a kit is provided, said kit comprising a matched set of fluorescent dyes comprising at least two different fluorescent dyes having the formula (I):

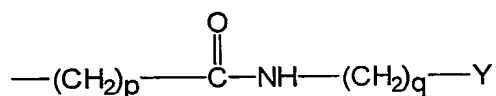


(I)

wherein n is different for each said dye and is 1, 2, or 3;

Z^1 and Z^2 independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system;

one of groups R^1 and R^2 is the group:



where Y is a target bonding group;

remaining group R^1 or R^2 is selected from $-(CH_2)_4-W$ or $-(CH_2)_r-H$;

group R^3 is hydrogen, except when either R^1 or R^2 is $-(CH_2)_r-H$, in which case R^3 is W ;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

r is an integer from 1 to 5;

and their salts;

characterised in that when n of two of said dyes differs by + 1, one of p , q and r of said two dyes differs by -1.

The invention is further illustrated by reference to the following examples and figures in which:

Figure 1 illustrates a typical workflow for differential analysis of protein samples using cysteine reactive dyes with 2D electrophoresis. Two different protein samples are prepared from cells, tissues or other biological fluids. The first and second protein samples may be any two samples in which it is desired to compare the protein content, for example, from a normal and diseased tissue, or from control versus drug treated/stimulated cells. The method may also be applied to one dimensional separation systems.

Figure 2 shows overlay images of proteins labelled with dye sets 1, 12 and 9 and separated by 2D electrophoresis with outlines of labelled protein spots to demonstrate the positional matching.

Examples

Example 1. Synthesis of Dyes

i) General Experimental Procedures

^1H NMR (δ_{H}) spectra were recorded on a Jeol JNM-LA300 FT NMR spectrometer. Chemical shifts are reported in δ (ppm). Samples were prepared as solutions in a suitable deuterated solvent such as d_4 -methanol. UV/VIS spectroscopy was conducted using the Unicam UV3 UV/VIS spectrometer. Trimethoxypropene was purchased from Karl Industries Inc., Ohio, USA. All other chemicals were purchased from Sigma-Aldrich Company Limited, Dorset, England.

N-BOC-Aminoethylmaleimide and N-BOC-aminopropylmaleimide were prepared according to the literature description in J. Org. Chem., (1995), 60, 5352-5355.

5 ii) Potassium 2,3,3-trimethylindolenine-5-sulphonic acid

Hydrazinobenzenesulphonic acid (20.0g) was dissolved in acetic acid (60ml) and 3-methyl-2-butanone (26.0g) added then heated at reflux for 3 hours. The desired compound was precipitated by cooling in the
10 fridge with scratching and the off white slurry was diluted with propan-2-ol and filtered (71%).

The 2,3,3-trimethyl-5-sulphonyl-indolenine (16.45g) was dissolved in methanol (160ml) with heating and a saturated solution of KOH in
15 propan-2-ol (100ml) was added. The solution changed to a yellow colour and a solid formed. The solution was cooled and the solid was filtered to form an off-white solid (15.9g, 98%). δ_H (300MHz, CD₃OD) 7.84 (m, 2H), 7.46 (d, 1H), 3.30 (s, 3H) and 1.35 (s, 6H).

20 iii) 1,2,3,3-Tetramethyl-5-sulphonyl-indolium iodide

Potassium 2,3,3-trimethylindolenine-5-sulphonic acid (1.0g, 3.61mmol) and iodomethane (0.25ml, 3.97mmol) were mixed with dichlorobenzene (10ml) under a nitrogen atmosphere. The solution was
25 heated at 100°C using a sand bath for 4 hours. A solid had begun to form but analysis by tlc (30% MeOH/70% DCM) showed product formation was not complete so an additional equivalent of iodomethane was added and the reaction heated for an additional 2 hours before cooling to room temperature. The solid was collected by filtration,
30 washed with dichlorobenzene, diethyl ether then dried *in vacuo* to afford

a purple solid (0.89g, 98%). δ_{H} (300MHz, CD_3OD) 8.06 (m, 1H), 7.94 (dd, 1H), 7.84 (m, 1H), 4.02 (s, 3H) and 1.61 (s, 6H).

iv) 1-Ethyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide

5

Potassium 2,3,3-trimethylindolenine-5-sulphonic acid (10.0g, 41.97mmol) and iodoethane (4.0ml, 50.35mmol) were mixed with dichlorobenzene (40ml) under a nitrogen atmosphere. The solution was heated at 120°C using a sand bath for 16 hours producing a purple solid.

10 The solid was collected by filtration then washed with dichlorobenzene, chloroform and ether to produce pale pink solid, (10.2g, 91%). δ_{H} (300MHz, CD_3OD) 7.98 (m, 3H), 4.55 (q, 2H), 1.56 (s, 6H) and 1.48 (t, 3H).

15 v) 1-Propyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide

Potassium 2,3,3-trimethylindolenine-5-sulphonic acid (1.0g, 3.61mmol) and iodopropane (0.4ml, 3.97mmol) were mixed with dichlorobenzene (10ml) under a nitrogen atmosphere. The solution was heated at 100°C using a sand bath for 20 hours producing a red-brown colour gelatinous solid. The solid was collected by filtration and then washed with dichlorobenzene, chloroform and diethyl ether to afford a pink solid, (472mg, 47%). δ_{H} (300MHz, CD_3OD) 7.91 (m, 3H), 4.50 (t, 2H), 2.01 (dt, 2H), 1.56 (s, 6H) and 1.13 (t, 3H).

25

vi) 1-Sulphobutyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide

2,3,3-Trimethylindolenine (2.0g, 12.6mmol) and 1,4-butanediol (1.7g, 12.6mmol) were mixed together and heated at 100°C for 6 hours.

30 The solution gradually turned red and after the 6 hours the reaction was cooled to room temperature. The solid was dispersed in diethyl ether and

filtered. The solid was collected and dried. δ_H (300MHz, CD_3OD) 7.91 (m, 1H), 7.75 (m, 1H), 7.42 (m, 2H), 4.58 (m, 2H), 2.93 (m, 2H), 2.18 (m, 2H), 1.94 (m, 2H) and 1.61 (s, 6H).

5 vii) 1-(5-Carboxypentyl)-2,3,3-trimethyl-5-sulphonyl-indolium iodide

2,3,3-Trimethylindolenine (6.4g, 40mmol) was dissolved in dichlorobenzene (25ml) and stirred until the solution was homogenous. To this was added 6-bromohexanoic acid (15.6g, 80mmol) and the
10 reaction heated to 110°C in a sand bath for 6.5 hours. The reaction was allowed to cool to room temperature where the sides of the flask were scratched then the flask was placed in the fridge for 1 hour. After this time, a beige solid had formed in the purple solution so the solid was collected by filtration then washed with dichlorobenzene and ether to
15 afford a beige solid (7.42g, 52%). δ_H (300MHz, CD_3OD) 7.91 (m, 1H), 7.78 (m, 1H), 7.62 (m, 2H), 4.52 (t, 2H), 2.38 (t, 2H), 2.04 (p, 2H), 1.88-2.45 (m, 4H) and 1.61 (s, 6H).

viii) 1-(4-Carboxybutyl)-2,3,3-trimethyl-5-sulphonyl-indolium iodide

20

2,3,3-Trimethylindolenine (11.3g, 40mmol) was dissolved in dichlorobenzene (30ml) and stirred until the solution was homogenous. To this was added 5-bromobutanoic acid (19.3g, 106.6mmol) and the reaction heated to 110°C in a sand bath for 6.5 hours. The reaction was
25 allowed to cool to room temperature where the sides of the flask were scratched then the flask was placed in the fridge for 1 hour. After this time a beige solid had formed in the purple solution so the solid was collected by filtration then washed with dichlorobenzene and ether to afford a beige solid (18.0g, 75%). δ_H (300MHz, CD_3OD) 7.90 (m, 1H),
30 7.78 (m, 1H), 7.62 (m, 2H), 4.59 (t, 2H), 2.41 (t, 2H), 2.04 (m, 2H), 1.76 (m, 2H) and 1.61 (s, 6H).

ix). Aminoethylmaleimide and Aminopropylmaleimide

To a solution of 4M hydrochloric acid in dioxane was added the
5 appropriate BOC-aminoalkylmaleimide and the reaction stirred at room
temperature for 30 minutes. After this time the reaction had deposited a
solid and the solvents were removed *in vacuo* to reveal a white fluffy
solid. The compound was used crude in subsequent steps.

10 x) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-
oxohexyl)-2-[(1E,3E)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2H-
indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound I)

1-Ethyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide (2.0g,
15 7.48mmol), N,N'-diphenylformamidine (1.5g, 7.48mmol) and
triethylorthoformate (1.1g, 7.48mmol) were dissolved in ethanol (10ml)
then heated at reflux (100°C) for 3 hours. A solid formed on the sides of
the reaction flask and UV/VIS showed a new peak at 408nm. Diethyl
ether was added and the precipitate and the solid collected by filtration,
20 washed with ether and dried in *vacuo* to afford a yellow/orange solid
(1.83g, 66%). UV/VIS (MeOH); absorption λ_{\max} = 408nm.

To a solution of the Cy3 half-dye (1.83g, 6.22mmol) in anhydrous
pyridine (10ml) was added acetic anhydride (1.0ml) and the reaction
25 stirred under a nitrogen atmosphere for 10 minutes. After this time 1-(5-
carboxypentyl)-2,3,3-trimethyl indolium bromide (2.2g, 6.22mmol) was
added and the reaction stirred at room temperature for 16 hours. The
progress of the reaction was monitored by tlc (20% MeOH/80% DCM).
The solvent was removed under reduced pressure and purified using flash
30 column chromatography (reversed phase silica: water-50% methanol

gradient) to yield 299mg of the desired Cy3 acid product (9%). UV/VIS (MeOH); absorption λ_{\max} = 550nm.

Cy3 acid (200mg, 0.36mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (80 μ l, 0.40mmol) and TSTU (120mg, 0.40mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (80 μ l, 0.40mmol) and aminoethylmaleimide (80mg, 0.40mmol) were added and the reaction allowed to stir for 2 hours. Thin layer chromatography (Tlc) showed conversion to a new product so the reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (98mg, 34%). UV/VIS (MeOH); absorption λ_{\max} = 550nm. δ_{H} (300MHz, CD₃OD) 8.56 (t, 1H), 7.92 (m, 2H), 7.62-7.38 (m, 5H), 6.72 (s, 2H), 6.50 (dd, 2H), 4.19 (m, 4H), 3.52 (m, 2H), 2.15 (t, 2H), 1.94-1.56 (m, 6H), 1.75 (s, 12H) and 1.44 (t, 3H).

xi) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1E,3E,5E)-5-(1,3,3-trimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)penta-1,3-dienyl]-3H-indolium (Compound II)

1,2,3,3-Tetramethyl-5-sulphonyl-indolium iodide (5.00g, 11.90mmol) was suspended in a mixture of acetic acid (40ml) and TFA (2ml, 18.0mmol) until all of the solid dissolved. 1,3,3-Trimethoxypropene (12.5ml, 95.0mmol) was added to the reaction and stirred at room temperature for 5 hours. The solution was pipetted into 500ml of diethyl ether and the precipitate collected by filtration (4.80g, contains salts).

To a solution of the Cy5 half-dye (4.80g, 14.9mmol) in methanol (40ml) was added potassium acetate (3.00g, 34.2mmol) and 1-(5-carboxypentyl)-2,3,3-trimethyl indolium bromide (3.00g, 16.4mmol). After stirring overnight, the solution was pipetted into diethyl ether (500ml) and the blue solid collected by filtration and dried *in vacuo*. Purification was achieved by flash column chromatography (reversed phase silica: water-50% methanol gradient) to yield 2.30g of desired product (28%). UV/VIS (MeOH); absorption λ_{\max} = 642nm.

Cy5 acid (200mg, 0.36mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (80 μ l, 0.40mmol) and TSTU (120mg, 0.40mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (80 μ l, 0.40mmol) and aminoethylmaleimide (80mg, 0.40mmol) were added and the reaction allowed to stir for 2 hours. Tlc showed conversion to a new product so the reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a blue residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (105mg, 43%). UV/VIS (MeOH); absorption λ_{\max} = 644nm. δ_{H} (300MHz, CD₃OD) 8.18 (dt, 2H), 7.82 (m, 2H), 7.56-7.24 (m, 5H), 6.80 (s, 2H), 6.64 (t, 1H), 6.45 (d, 1H), 6.24 (d, 1H), 4.17 (t, 2H), 3.60 (s, 3H), 3.40 (t, 2H), 2.09 (t, 2H), 1.89-1.48 (m, 6H) and 1.79 (s, 12H).

xii) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(1-propyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound III)

1-(5-Carboxypentyl)-2,3,3-trimethyl-5-sulphonyl-indolium (6.0g, 16.94mmol) and N,N'-diphenylformamidine (6.63g, 33.87mmol) were

dissolved in acetic acid (20ml) then heated at reflux (120°C) for 4 hours. The reaction was allowed to cool to room temperature then the solvents removed *in vacuo*. The orange oil was dissolved in chloroform and washed with water then dried with magnesium sulphate and concentrated to an orange oil of the Cy3 half-dye. This oil was used without further purification.

To a solution of the Cy3 half-dye (1.00g, 2.18mmol) in anhydrous pyridine (10ml) was added acetic anhydride (1.0ml) and the reaction stirred under a nitrogen atmosphere for 10 minutes. After this, 1-propyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide (0.92g, 3.28mmol) was added and the reaction stirred at room temperature for 16 hours. The progress of the reaction was monitored by tlc (20% MeOH/80% DCM). The solvent was removed under reduced pressure and purified using flash column chromatography (silica: DCM-40% methanol gradient) to yield 241mg of the desired Cy3 acid product (21%). UV/VIS (MeOH); absorption $\lambda_{\max} = 550\text{nm}$.

Cy3 acid (100mg, 0.18mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (30 μl , 0.20mmol) and TSTU (54mg, 0.20mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (30 μl , 0.20mmol) and aminoethylmaleimide (40mg, 0.20mmol) were added and the reaction allowed to stir for 2 hours. Tlc showed conversion to a new product so the reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (41mg, 34%). UV/VIS (MeOH); absorption $\lambda_{\max} = 550\text{nm}$. δ_{H} (300MHz, CD_3OD) 8.52 (t, 1H), 7.90 (m,

2H), 7.60-7.39 (m, 5H), 6.78 (s, 2H), 6.52 (dd, 2H), 4.22 (m, 4H), 3.56 (m, 2H), 2.13 (t, 2H), 1.97-1.56 (m, 8H), 1.75 (s, 12H) and 1.09 (t, 3H).

xiii) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound IV)

1-(5-Carboxypentyl)-2,3,3-trimethyl-5-sulphonyl-indolium (4.0g, 14.59mmol) and malonaldehyde bis(phenylimine) hydrochloride (7.55g, 29.18mmol) were dissolved in acetic acid (20ml) then heated at reflux (120°C) for 4 hours. The reaction was allowed to cool to room temperature then the solvents removed *in vacuo*. The red oil was dissolved in chloroform and washed with water then dried with magnesium sulphate and concentrated to a red oil of the Cy5 half-dye. This oil was used without further purification.

To a solution of the Cy5 half-dye (1.00g, 2.07mmol) in anhydrous pyridine (10ml) was added acetic anhydride (1.0ml) and the reaction stirred under a nitrogen atmosphere for 10 minutes. To this 1-ethyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide (0.59g, 2.28mmol) was added and the reaction stirred at room temperature for 16 hours. The progress of the reaction was monitored by tlc (20% MeOH/80% DCM). The solvent was removed under reduced pressure and purified using flash column chromatography (silica: DCM-40% methanol gradient) to yield 365mg of the desired Cy5 acid product (31%). UV/VIS (MeOH); absorption $\lambda_{\max} = 644\text{nm}$.

Cy5 acid (52mg, 0.09mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (10 μ l, 0.10mmol) and TSTU (28mg, 0.10mmol) were added and the

reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (10 μ l, 0.10mmol) and aminoethylmaleimide (18mg, 0.10mmol) were added and the reaction allowed to stir for 2 hours. Tlc showed conversion to a new product so the reaction was diluted with diethyl ether (30ml) and the solvents decanted to leave a blue residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (29mg, 47%). UV/VIS (MeOH); absorption λ_{max} = 644nm. δ_{H} (300MHz, CD₃OD) 8.38 (m, 2H), 7.88 (m, 2H), 7.55-7.24 (m, 5H), 6.80 (s, 2H), 6.71 (t, 1H), 6.42 (d, 1H), 6.24 (d, 1H), 4.11 (m, 4H), 3.54 (t, 2H), 2.12 (t, 2H), 1.92-1.36 (m, 9H) and 1.81 (s, 12H).

xiv) 1-(5-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1E,3E,5E)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)penta-1,3-dienyl]-3H-indolium (Compound V)

1-(4-Carboxybutyl)-2,3,3-trimethyl indolium bromide (2.00 g, 5.88 mmol) and malonaldehyde bis(phenylimine) hydrochloride (2.28g, 8.82mmol) were dissolved in acetic acid (30ml) then heated at 120°C for 6 hours. The reaction was then allowed to cool before the acetic acid was removed *in vacuo* afford a mobile oil. This was dissolved in chloroform and washed with water, dried with magnesium sulphate, filtered and concentrated *in vacuo* to afford a more viscous oil. Most of the unreacted malonaldehyde bis(phenylimine) remained at the interface of the chloroform/ water. The compound was purified using flash column chromatography (dichloromethane-30% methanol gradient) to afford a red solid (120mg, 5%).

To a solution of the Cy5 half-dye (120mg, 0.31mmol) in anhydrous pyridine (3ml) was added acetic anhydride (0.5ml) and the reaction stirred under a nitrogen atmosphere for 10 minutes. After this time, the 1-ethyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide (83mg, 0.31mmol) was added and the reaction stirred at room temperature overnight. The solvent was removed under reduced pressure and purified using flash column chromatography (silica: dichloromethane-40% methanol gradient) to yield 84 mg of desired product (48%). UV/VIS (MeOH); absorption λ_{\max} = 644nm.

Cy5 (84mg, 0.15mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (26 μ l, 0.15mmol) and TSTU (45mg, 0.15mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (DIPEA (26 μ l, 0.15mmol) and then aminoethylmaleimide (52mg, 0.30mmol) was added and the reaction allowed to stir for an additional 2 hours. Tlc showed conversion to a new product, so the reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (30 mg, 30%). UV/VIS (MeOH); absorption λ_{\max} = 644nm. δ_{H} (300MHz, CD₃OD) 8.15 (q, 2H), 7.80 (m, 2H), 7.52-7.31 (m, 5H), 6.75 (s, 2H), 6.64 (t, 1H), 6.46 (d, 1H), 6.32 (d, 1H), 4.27 (m, 4H), 4.15 (t, 2H), 2.09 (t, 2H), 1.79 (s, 6H), 1.65 (s, 6H) 2.56-2.44 (m, 4H) and 1.12 (t, 3H).

xv) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound VI)

5 1-(5-Carboxypentyl)-2,3,3-trimethyl indolium bromide (500mg, 1.41mmol), N,N'-diphenylformamidine (278mg, 1.41mmol) and 1-sulphobutyl-2,3,3-trimethylindolium iodide (418mg, 1.41mmol) were dissolved in anhydrous pyridine (5ml) and stirred at room temperature. Acetic anhydride (0.4ml) was then added and the reaction stirred
10 overnight at ambient temperature. The pyridine was removed *in vacuo* and the magenta oil purified by flash column chromatography (dichloromethane-10% methanol gradient) to afford a pink solid (67mg, 8%).

15 The Cy3 acid (50mg, 0.09mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (20 μ l, 0.09mmol) and TSTU (59mg, 0.18mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second
20 equivalent of DIPEA (20 μ l, 0.09mmol) and then the aminoethylmaleimide (32mg, 0.18mmol) was added and the reaction allowed to stir for an additional 2 hours. The reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the
25 desired maleimide product (24mg, 43%). UV/VIS (MeOH); absorption λ_{\max} = 550nm. δ_{H} (300MHz, CD₃OD) 8.49 (t, 1H), 7.52-7.35 (m, 8H), 6.72 (s, 2H), 6.46 (dd, 2H), 4.09 (m, 4H), 3.57 (m, 2H), 2.85 (m, 2H), 2.15 (t, 2H), 1.94-1.43 (m, 10H) and 1.85 (s, 12H).

xvi) 1-(5-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1E,3E,5E)-5-(3,3-dimethyl-(1-sulpho-butyl)-1,3-dihydro-2H-indol-2-ylidene)penta-1,3-dienyl]-3H-indolium (Compound VII)

5

1-(4-Carboxybutyl)-2,3,3-trimethyl indolium bromide (1.0g, 2.94mmol) and malonaldehyde bisphenylimine hydrochloride (760mg, 2.94mmol) were dissolved in acetic acid (10ml) then heated at 120°C for 18 hours. The reaction was then allowed to cool before the acetic acid was removed *in vacuo* afford a mobile red oil. This was dissolved in chloroform and washed with water, dried with magnesium sulphate, filtered and concentrated *in vacuo* to afford a more viscous red/brown oil. The compound was purified using flash column chromatography (dichloromethane-20% methanol gradient) to afford an orange solid (1.43g, 55%).

15

To a solution of the Cy5 half-dye (120mg, 0.26mmol) in anhydrous pyridine (5ml) was added acetic anhydride (0.5ml) and the reaction stirred under a nitrogen atmosphere for 10 minutes. After this time 1-sulphobutyl-2,3,3-trimethylindolium iodide (83mg, 0.28mmol) was added and the reaction stirred at room temperature for 2 hours. The solvent was removed under reduced pressure and purified using flash column chromatography (silica: dichloromethane-20% methanol gradient) to yield 78 mg of desired product (51%).

25

The Cy5 acid (50mg, 0.09mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (15µl, 0.10mmol) and TSTU (28mg, 0.10mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (15µl, 0.10mmol) and then aminoethylmaleimide

30

(30mg, 0.17mmol) was added and the reaction allowed to stir for an additional 2 hours. Tlc analysis showed conversion to a new product.

The reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a blue residue. Flash column chromatography (silica:

5 DCM-40% methanol gradient) afforded the desired maleimide product (30mg, 30%). UV/VIS (MeOH); absorption λ_{max} = 644nm. δ_{H} (300MHz, CD₃OD) 8.19 (t, 2H), 7.52-7.42 (m, 8H), 6.71 (s, 2H), 6.64 (t, 1H), 6.36 (dt, 2H), 4.18 (m, 4H), 3.55 (m, 2H), 2.89 (m, 2H), 2.15 (t, 2H), 1.79 (s, 12H) and 2.01-1.36 (t, 8H).

10

xvii) 1-(6-{[3-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound VIII)

15 1-(5-Carboxypentyl)-2,3,3-trimethyl indolium bromide (354mg, 1.00mmol), N,N'-diphenylformamidine (196mg, 1.00mmol) and 1-ethyl-2,3,3-trimethyl-5-sulphonyl-indolium iodide (267mg, 1.00mmol) were dissolved in anhydrous pyridine (15ml) and stirred at room temperature. Acetic anhydride (0.5ml) was then added and the reaction stirred
20 overnight at ambient temperature. The pyridine was removed *in vacuo* and the magenta oil purified by flash column chromatography (dichloromethane-40% methanol gradient) to afford a pink solid (30mg, 6%).

25 The Cy3 acid (30mg, 0.05mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (10 μ l, 0.05mmol) and TSTU (2mg, 0.05mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second
30 equivalent of DIPEA (10 μ l, 0.05mmol) and then the aminoethylmaleimide (9mg, 0.05mmol) was added and the reaction allowed to stir for an

additional 2 hours. The reaction was diluted with diethyl ether (20ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (15mg, 44%). UV/VIS (MeOH); absorption λ_{\max} = 550nm. δ_{H} (300MHz, CD_3OD) 8.53 (t, 1H), 7.95 (m, 2H), 7.59-7.41 (m, 5H), 6.78 (s, 2H), 6.50 (t, 2H), 4.17 (m, 4H), 3.50 (m, 2H), 3.14 (t, 2H), 2.14 (t, 2H), 1.98-1.28 (m, 11H) and 1.79 (s, 12H).

xviii) 1-(6-{[3-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound IX)

1-(5-Carboxypentyl)-2,3,3-trimethyl indolium bromide (6.0g, 16.95mmol) and N,N'-diphenylformamidine (6.63g, 33.87mmol) were dissolved in acetic acid (20ml) then heated at 120°C for 5 hours. The acetic acid was removed *in vacuo* and the residue purified using column chromatography (silica: dichloromethane-20% methanol gradient) yellow/orange solid (1.34g, 21%).

To a solution of the Cy3 half-dye (250mg, 0.55mmol) in anhydrous pyridine (5ml) was added acetic anhydride (0.5ml) and the reaction stirred under a nitrogen atmosphere for 10 minutes. After this time 1-sulphobutyl-2,3,3-trimethylindolium iodide (161mg, 0.55mmol) was added and the reaction stirred at room temperature for 22 hours. The solvent was removed under reduced pressure and purified using flash column chromatography (silica: dichloromethane-20% methanol gradient) to yield 142mg of the desired Cy3 acid product (45%).

Cy3 acid (50mg, 0.09mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (15 μ l, 0.10mmol) and TSTU (28mg, 0.10mmol) were added and the

reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (15µl, 0.10mmol) and aminopropylmaleimide (18mg, 0.10mmol) were added and the reaction allowed to stir for 2 hours. The reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (15mg, 24%). δ_H (300MHz, CD₃OD) 8.49 (t, 1H), 7.78-7.22 (m, 8H), 6.72 (s, 2H), 6.50 (t, 2H), 4.12 (m, 4H), 3.47 (m, 2H), 3.10 (m, 2H), 2.85 (m, 2H), 2.15 (t, 2H), 2.01-1.24 (m, 12H) and 1.85 (s, 12H).

xix) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1E,3E,5E)-5-(3,3-dimethyl-(1-sulpho-butyl)-1,3-dihydro-2H-indol-2-ylidene)penta-1,3-dienyl]-3H-indolium (Compound X)

1-(5-Carboxypentyl)-2,3,3-trimethyl indolium bromide (353mg, 1.00mmol), malonaldehyde bisphenylimine hydrochloride (259mg, 1.00mmol) and 1-sulphobutyl-2,3,3-trimethylindolium iodide (295mg, 1.00mmol) were dissolved in anhydrous pyridine (5ml) and stirred at room temperature. Acetic anhydride (0.4ml) was then added and the reaction stirred overnight at ambient temperature. The pyridine was removed *in vacuo* and the blue oil purified by flash column chromatography (dichloromethane-10% methanol gradient) to afford a blue solid (54mg, 8%).

The Cy5 acid (49mg, 0.08mmol) was dissolved in anhydrous DMF under a nitrogen atmosphere then stirred at room temperature. DIPEA (12µl, 0.09mmol) and TSTU (27mg, 0.09mmol) were added and the reaction stirred for 2 hours until deemed complete to the NHS ester by tlc (20% MeOH/80% DCM). The NHS ester was treated with a second equivalent of DIPEA (12µl, 0.09mmol) and then aminoethylmaleimide

(28mg, 0.16mmol) was added and the reaction allowed to stir for an additional 2 hours. Tlc showed conversion to a new product so the reaction was diluted with diethyl ether (50ml) and the solvents decanted to leave a pink residue. Flash column chromatography (silica: DCM-40% methanol gradient) afforded the desired maleimide product (24mg, 45%). UV/VIS (MeOH); absorption λ_{\max} = 644nm. δ_{H} (300MHz, CD₃OD) 8.21 (t, 2H), 7.54-7.26 (m, 8H), 6.80 (s, 2H), 6.64 (t, 1H), 6.38 (dd, 2H), 4.18 (m, 4H), 3.58 (m, 2H), 2.89 (m, 2H), 2.15 (t, 2H), 1.76 (s, 12H) and 2.06-1.39 (m, 10H).

xx) 1-(6-{[2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(1,3,3-trimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium (Compound XI)

By similar methods, 1,2,3,3-tetramethyl-5-sulphonyl-indolium iodide and 1-(5-carboxypentyl)-2,3,3-trimethyl indolium bromide were reacted together with N,N'-diphenylformamidine to form the Cy3 acid, which when activated to the N-hydroxysuccinimide ester and treated with aminoethylmaleimide, was converted to 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1E,3E)-3-(1,3,3-trimethyl-5-sulpho-1,3-dihydro-2H-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3H-indolium.

Example 2

2.1 Protein Isolation and Labelling

Initial experiments were performed on *E. coli* samples. *E. coli* strain ER1647 (Amersham Biosciences, Buckinghamshire, UK) was grown in glucose rich MOPS media at 37°C overnight, followed by harvesting by centrifugation for 10 minutes at 4°C at 12,000 x g. The cell pellet was

washed twice with wash buffer (10mM Tris pH 8.0, 0.5mM magnesium acetate). Cells were then resuspended in lysis buffer (8M urea, 4% w/v CHAPS, 40mM Tris pH8.0) and lysed by sonication (3 x 10 second pulses on ice). The protein concentration of the *E.coli* lysate was determined using the Bio-Rad Dc Protein Assay as described by the manufacturer (Bio-Rad, Hertfordshire, UK).

Before use, the cyanine dyes were reconstituted in anhydrous DMF (Aldrich catalogue code 22,705-6) to give a final concentration of 10mM (10nmol/ μ l). The dye was vortexed briefly after addition of DMF to ensure the dye was completely dissolved. The reconstituted dye was stored at -20°C and used within 2 weeks.

Disulphide bonds in 50 μ g protein in lysis buffer were reduced by the addition of 10nmol TCEP [tris-(2-carboxyethyl)phosphine] followed by incubation at 37°C for 1 hour. Following reduction, 20nmol reconstituted dye was added and incubated for 30 minutes at 37°C . The reaction was stopped with an equal volume of 2x sample buffer (Lysis buffer plus 20mg / ml DTT and 4%(v/v) Pharmalytes 3-10). The samples were stored briefly on ice prior to first dimension separation or frozen protected from light for longer storage.

2.2 Preparation of protein samples for separation.

Equal amounts of protein samples labelled with Cy3 and Cy5 compounds I – XI were mixed to give the dye pair sets shown in Table 2.

Table 2

<u>Dye pair set</u>	<u>Compound</u>	<u>Dye pair set</u>	<u>Compound</u>
1	I II	7	VIII II
2	III IV	8	IX VII
3	I V	9	III II
4	VI VII	10	VI X
5	VIII IV	11	I IV
6	IX X	12	XI II

2.3 Protein Separation by 2D Electrophoresis

5 2-D electrophoresis was performed using standard Amersham Biosciences 2D PAGE equipment and PlusOne™ reagents (Buckinghamshire, UK). Immobiline DryStrips (pH3-10 NL or pH 4-7 NL, 24cm) were rehydrated overnight in 450µl rehydration buffer (8M urea, 4% w/v CHAPS, 1% Pharmalytes (pH 3-10), 2mg/ml DTT) overlaid with

10 2.5ml DryStrip Cover Fluid, in an Immobiline DryStrip Reswelling Tray. Strips were focused using the Multiphor isoelectric focusing system. Prior to 2nd dimension PAGE, each strip was equilibrated with 10ml equilibration buffer A (8M urea, 100mM Tris-HCl pH6.8, 30% v/v glycerol, 1% w/v SDS, 5mg/ml DTT) on a rocking table for 10 minutes,

15 followed by 10ml equilibration buffer B (8M urea, 100mM Tris-HCl pH6.8, 30% v/v glycerol, 1% w/v SDS, 45mg/ml iodoacetamide) for a further 10 minutes. The strips were then loaded and run on 12% isochratic Laemmli SDS-PAGE gels.

2.4 Fluorescence Gel Imaging

Labelled proteins were visualised using the 2920 MasterImager
(Amersham Biosciences, Buckinghamshire, UK) with the following
settings:

	<u>Excitation</u>	<u>Emission</u>
Cy3	540nm (25nmBP)	590nm (35nmBP)
Cy5	620nm (30nmBP)	680nm (30nmBP)

Exposure times were optimised for individual experiments to give a
maximum pixel value on the image of 50,000 to avoid saturation of the
signal. Data from 2D-Master was exported as TIF files into Paintshop
Pro™ to generate colour overlays for visual inspection. For detailed
quantitative analysis of dye matching, gel images were exported into 2D
Image Master software programme.

2.5 Image Analysis

Gel analysis in this study was performed using 2D Image Master
(Amersham Biosciences, Buckinghamshire, UK) a 2-D analysis software
platform. Following spot detection, the centre of each spot was used to
generate pixel co-ordinates for each of the Cy3 and Cy5 images. The
extent of migration matching was determined by comparing the co-
ordinate positions of the centres of the Cy3 and Cy5 spots. This can be
used to describe positional matching in terms of the number of matched
spots within ± 2 pixels of each other in both the x (pI) and y (mass) co-
ordinates.

As both dyes have been used to label the same sample the same proteins are present in each labelled sample and thus if the dyes are matched for equivalent migration all of the spots should exactly overlay between the 2 images. The % of spots within +/- 2 pixels was
5 measured to determine the overall matching.

Example 3 Difference Gel Electrophoresis of Dyes Matched for Saturation Labelling on Cysteine Residues by Variation of the Alkyl Chain Length (r)

10 Dye structure effects were tested by labelling the same complex protein sample (E.coli lysate) with each dye, mixing the Cy3 and Cy5 labelled samples and separating by 2D electrophoresis. Following fluorescence scanning, the images were overlaid and analysed as
15 described in the above example.

E.coli lysate was labelled with Cy3 compounds and Cy5 compounds as described. The labelled samples were mixed prior to separation to give dye sets 1, 2, 9, 11 and 12 to show the effect of
20 varying the length of the alkyl chain by the addition of CH₂ units. Table 3 shows the quantitative positional data following analysis of the overlaid images of protein labelled with dye sets 1, 2, 9, 11 and 12.

25

30

Table 3: Effect of Varying Alkyl Chain Length (r) on Positional Matching

Dye Set	Value of n	p	q	r	Matched Spots within 2 Pixels (%)	
					X (pl)	Y (mass)
1	1 (Cy3)	5	2	2	82.7	88.8
	2 (Cy5)	5	2	1		
2	1 (Cy3)	5	2	3	79.7	77.7
	2 (Cy5)	5	2	2		
9	1 (Cy3)	5	2	3	76	76.3
	2 (Cy5)	5	2	1		
11	1 (Cy3)	5	2	2	86.1	44.3
	2 (Cy5)	5	2	2		
12	1 (Cy3)	5	2	1	59.7	58.9
	2 (Cy5)	5	2	1		

Thus, when the Cy3 and Cy5 have equal alkyl chains (e.g. dye sets 11 and 12) the overall mass of Cy3 is decreased relative to Cy5. This results in poor positional matching in the mass dimension. When there is a two carbon unit difference in the alkyl chain length (e.g. dye set 9), mass matching improves. Optimum matching for both mass and pl is obtained when there is a single carbon atom difference between the dyes (e.g. dye set 1 and 2). The preferred dye set is dye set 1. This shows that migration matching using a single carbon unit difference in the linker r gives good positional matching of differentially labelled proteins on 2D electrophoresis.

Figure 2 shows overlay images of proteins labelled with dye sets 1, 12 and 9 and separated by 2D electrophoresis with outlines of labelled protein spots to demonstrate the positional matching. The overlays are

taken from a portion of a 2D electrophoresis gel of E.coli lysate labelled with Cy3 or Cy5 showing the ~20 - 30kDa mass range and ~ 5.5 - 6.0 pI range. The circles representing the outline positions of the protein spots for both the Cy3 and Cy5 labelled proteins (arrowed) were

5 determined using the 2D ImageMaster software. Dye set 12 gives poor mass matching as the Cy5 labelled proteins run below the Cy3 labelled proteins (example arrowed). Dye set 9 also shows poor mass matching, in this case with the Cy5 labelled proteins running above the Cy3 labelled proteins (example arrowed). Preferred dye set 1 gives the optimal

10 positional matching.

Example 4 Difference Gel Electrophoresis of Dyes Matched for Saturation Labelling on Cysteine Residues by Variation of the Length of the Linker to the Indole Nucleus (p)

15

E.coli lysate was labelled with Cy3 and Cy5 compounds and samples mixed to give dye sets 1 and 3 as described. Table 4 shows the quantitative positional data following analysis of the overlaid images of protein labelled with these dye sets.

20

Table 4: Effect of Varying Linker Length (p) on Positional Matching

Dye Set	Value of n	p	q	r	Matched Spots within 2 Pixels (%)	
					X (pI)	Y (mass)
1	1 (Cy3)	5	2	2	87.3	86.4
	2 (Cy5)	5	2	1		
3	1 (Cy3)	5	2	2	88.1	85.2
	2 (Cy5)	4	2	2		

This shows that migration matching using a single carbon unit difference in the linker p gives good positional matching of differentially labelled proteins on 2D electrophoresis.

5 Example 5 Difference Gel Electrophoresis of Dyes Matched for Saturation Labelling on Cysteine Residues by Variation of the Length of the Linker to the Maleimide Group (q)

10 Variation of the linker q was achieved by increasing the length of the linker by a single CH₂ unit from an aminoethyl to aminopropyl maleimide. E.coli lysate was labelled with Cy3 and Cy5 compounds and samples mixed to give dye sets 1, 5 and 7 as described. Table 5 shows the quantitative positional data following analysis of the overlaid images of protein labelled with these dye sets.

15

Table 5: Effect of Varying Maleimide Linker Length (q) on Positional Matching

Dye Set	Value of n	p	q	r	Matched Spots within 2 Pixels (%)	
					X (pl)	Y (mass)
1	1 (Cy3)	5	2	2	85.7	88.0
	2 (Cy5)	5	2	1		
5	1 (Cy3)	5	3	2	64.2	80.7
	2 (Cy5)	5	2	2		
7	1 (Cy3)	5	3	2	67.6	71.8
	2 (Cy5)	5	2	1		

20 Thus, the aminoethyl linker (dye set 1) has improved pl matching compared to the aminopropyl linker (dyes sets 5 and 7) possibly due to the β -effect. When there is a two carbon unit difference between the

dyes the mass matching also decreases (dye set 7). When there is a single carbon unit difference mass matching improves. Thus, an aminoethyl maleimide linker ($q = 2$) is preferred for pl matching.

5 **Example 6 Difference Gel Electrophoresis of Dyes Matched for Saturation Labelling on Cysteine Residues by Variation of the Position of the Sulphonate Group**

10 Sulphonated cyanine dyes generally have the sulphonate group directly attached to the indole ring as in dye set 1. Variation of the sulphonate position was achieved by attaching the sulphonate to the ring via a butyl chain linker. E.coli lysate was labelled with Cy3 and Cy5 compounds and samples mixed to give dye sets 1, 4, 6, 8 and 10 as described. Table 6 shows the quantitative positional data following
15 analysis of the overlaid images of protein labelled with these dye sets.

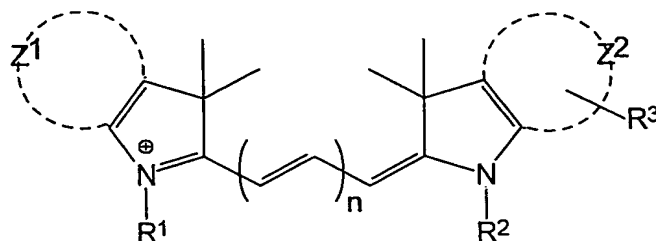
Table 6: Effect of Varying Sulphonate Position on Matching

Dye Set	Sulphonate Position	Value of n	p	q	r	Matched Spots within 2 Pixels (%)	
						X (pl)	Y (mass)
6	Pendant	1 (Cy3)	5	3	4	86.3	82.6
	Pendant	2 (Cy5)	5	2	4		
4	Pendant	1 (Cy3)	5	2	4	78.3	79.7
	Pendant	2 (Cy5)	4	2	4		
8	Pendant	1 (Cy3)	5	3	4	73.3	63.7
	Pendant	2 (Cy5)	4	2	4		
10	Pendant	1 (Cy3)	5	2	4	82.4	49.8
	Pendant	2 (Cy5)	5	2	4		

Comparison with previous data shows that the removal of the sulphonate from the ring system to a pendant butyl chain does appear to change the pl matching. The matching in the mass dimension is no better than that achieved with dye set 1. Where there is no compensation for the difference n between the dyes the migration matching decreases (dye set 10). When there is a two carbon unit difference (dye set 8), mass matching is decreased compared to dye set 1. When there is a single carbon unit difference (dye sets 4 and 6) matching is improved. The best combination with a pendant butyl sulphonate is to modify the linker q to the maleimide (dye set 6).

Claims

1. A matched set of fluorescent dyes comprising at least two different fluorescent dyes of formula (I):

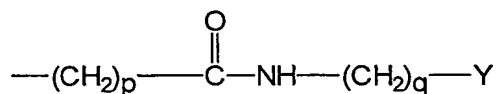


(I)

wherein n is different for each said dye and is 1, 2, or 3;

Z¹ and Z² independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system;

one of groups R¹ and R² is the group:



where Y is a target bonding group;

remaining group R¹ or R² is selected from $-(\text{CH}_2)_4\text{W}$ or $-(\text{CH}_2)_r\text{H}$;

group R³ is hydrogen, except when either R¹ or R² is $-(\text{CH}_2)_r\text{H}$, in which case R³ is W;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

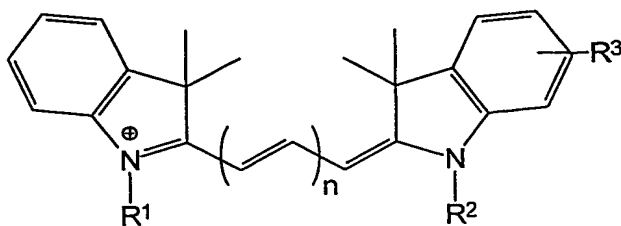
q is selected to be 2 or 3; and

r is an integer from 1 to 5;

and their salts;

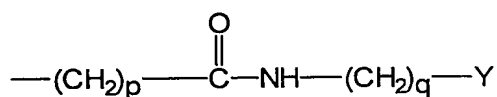
characterised in that when n of two of said dyes differs by + 1, one of p, q and r of said two dyes differs by -1.

2. A matched set of fluorescent dyes comprising at least two different fluorescent dyes of formula (II):



(II)

10 wherein n is different for each said dye and is 1, 2, or 3;
one of groups R^1 and R^2 is the group:



where Y is a target bonding group;

remaining group R^1 or R^2 is selected from $-(CH_2)_4-W$ or $-(CH_2)_r-H$;

group R^3 is hydrogen, except when either R^1 or R^2 is $-(CH_2)_r-H$, in which case R^3 is W ;

20 W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

r is an integer from 1 to 5;

and their salts;

25 characterised in that when n of two of said dyes differs by + 1, one of p , q and r of said two dyes differs by -1.

3. A matched set according to claim 1 or claim 2 comprising at least two different fluorescent dyes according to formula (I) or (II) in which:

30 n is selected to be 1 or 2;

p is selected to be 4 or 5;

q is selected to be 2 or 3; and
r is selected to be 1, 2 or 3.

4. A matched set according to any of claims 1 to 3 wherein said
5 target bonding group Y in each dye of the set of dyes is the same and is
selected from a maleimido group and an iodoacetamido group.

5. A matched set according to claim 4 wherein in each said dye Y is a
maleimido group.

10

6. A matched set according to any of claims 1 to 5 wherein said salts
are selected from K^+ , Na^+ , NH_4^+ , R_3NH^+ and R_4N^+ where R is C_1 to C_4
alkyl.

15 7. A matched set of dyes according to any of claims 1 to 6 selected
from:

Set 1

20 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-
[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-
ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound I); and
1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-
3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1,3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-
25 indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound II);

Set 2

30 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-
[(1*E*,3*E*)-3-(1-propyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-
ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound III); and

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound IV);

5 Set 3

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound I); and

10 1-(5-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound V);

Set 4

15

1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound VI); and

20 1-(5-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxopentyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(3,3-dimethyl-(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound VII).

Set 5

25 1-(6-{[3-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-2-[(1*E*,3*E*)-3-(1-ethyl-3,3-dimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound VIII); and

30 1-(6-{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(1-ethyl-3,3-trimethyl-5-sulpho-1,3-dihydro-2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound IV); and

Set 6

1-(6-{{[3-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)propyl]amino}-6-oxohexyl)-
2-[(1*E*,3*E*)-3-(3,3-dimethyl(1-sulpho-butyl)-1,3-dihydro-2*H*-indol-2-
5 ylidene)prop-1-enyl]-3,3-dimethyl-3*H*-indolium (Compound IX); and
1-(6-{{[2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)ethyl]amino}-6-oxohexyl)-
3,3-dimethyl-2-[(1*E*,3*E*,5*E*)-5-(3,3-dimethyl-(1-sulpho-butyl)-1,3-dihydro-
2*H*-indol-2-ylidene)penta-1,3-dienyl]-3*H*-indolium (Compound X).

10 8. A method for labelling a mixture of proteins in a sample wherein
each of said proteins contains one or more cysteine residues, said method
comprising:

- i) adding to an aqueous liquid containing said sample a fluorescent
dye wherein said dye contains a target bonding group that is covalently
15 reactive with said proteins; and
ii) reacting said dye with said proteins so that said dye labels said
proteins;
characterised in that all available cysteine residues in said proteins are
labelled with said dye.

20

9. A method according to claim 8 wherein said fluorescent dye is a
cyanine dye.

10. A method according to claim 9 wherein said cyanine dye contains a
25 sulphonic acid or sulphonate group.

11. A method according to any of claims 8 to 10 wherein said target
bonding group is selected from a maleimido group and an iodoacetamido
group.

30

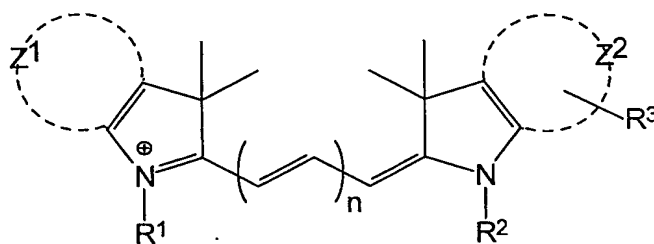
12. A method according to claim 8 further comprising prior to step i), the step of treating the protein with a reductant.

13. A method according to claim 8 wherein said dye is used in a range of 5 to 200nmol of dye per 50 μ g of protein.

14. A method according to claim 8 wherein said labelling is performed at a pH in the range from 6.0 to 9.0.

15. A method for labelling one or more proteins in a sample, the method comprising:

i) adding to a liquid sample containing said one or more proteins a fluorescent dye of formula (I):

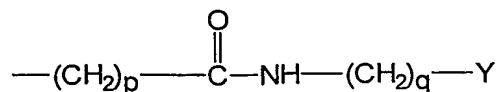


(I)

wherein n is different for each said dye and is 1, 2, or 3;

Z¹ and Z² independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system;

one of groups R¹ and R² is the group:



where Y is a target bonding group;

remaining group R¹ or R² is selected from -(CH₂)₄-W or -(CH₂)_r-H;

group R³ is hydrogen, except when either R¹ or R² is -(CH₂)_r-H, in which case R³ is W;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

r is an integer from 1 to 5;

5 and their salts;

characterised in that when n of two of said dyes differs by +1, one of p, q and r of said two dyes differs by -1; and

ii) incubating said dye with said sample under conditions suitable for labelling said one or more proteins.

10

16. A method according to claim 15 wherein each of Z^1 and Z^2 represents the carbon atoms necessary to complete a phenyl ring system.

17. A method according to claim 15 or claim 16 wherein:

15 n is selected to be 1 or 2;

p is selected to be 4 or 5;

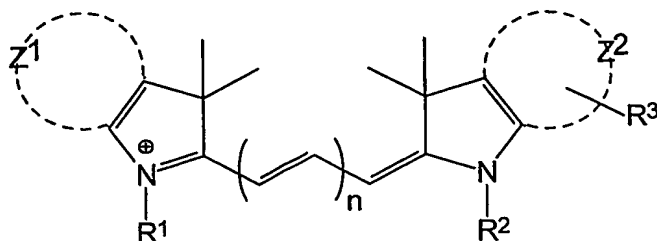
q is selected to be 2 or 3; and

r is selected to be 1, 2 or 3.

20 18. A method according to any of claims 15 to 17 wherein said target bonding group Y is selected from a maleimido group and an iodoacetamido group.

19. A kit comprising a matched set of fluorescent dyes comprising at

25 least two different fluorescent dyes having the formula (I):

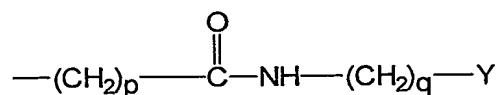


(I)

wherein n is different for each said dye and is 1, 2, or 3;

Z^1 and Z^2 independently represent the carbon atoms necessary to complete a phenyl or naphthyl ring system;

one of groups R^1 and R^2 is the group:



where Y is a target bonding group;

remaining group R^1 or R^2 is selected from $-(CH_2)_4-W$ or $-(CH_2)_r-H$;

group R^3 is hydrogen, except when either R^1 or R^2 is $-(CH_2)_r-H$, in which case R^3 is W ;

W is selected from sulphonic acid and sulphonate;

p is an integer from 3 to 6;

q is selected to be 2 or 3; and

r is an integer from 1 to 5;

and their salts;

characterised in that when n of two of said dyes differs by + 1, one of p ,

q and r of said two dyes differs by -1.

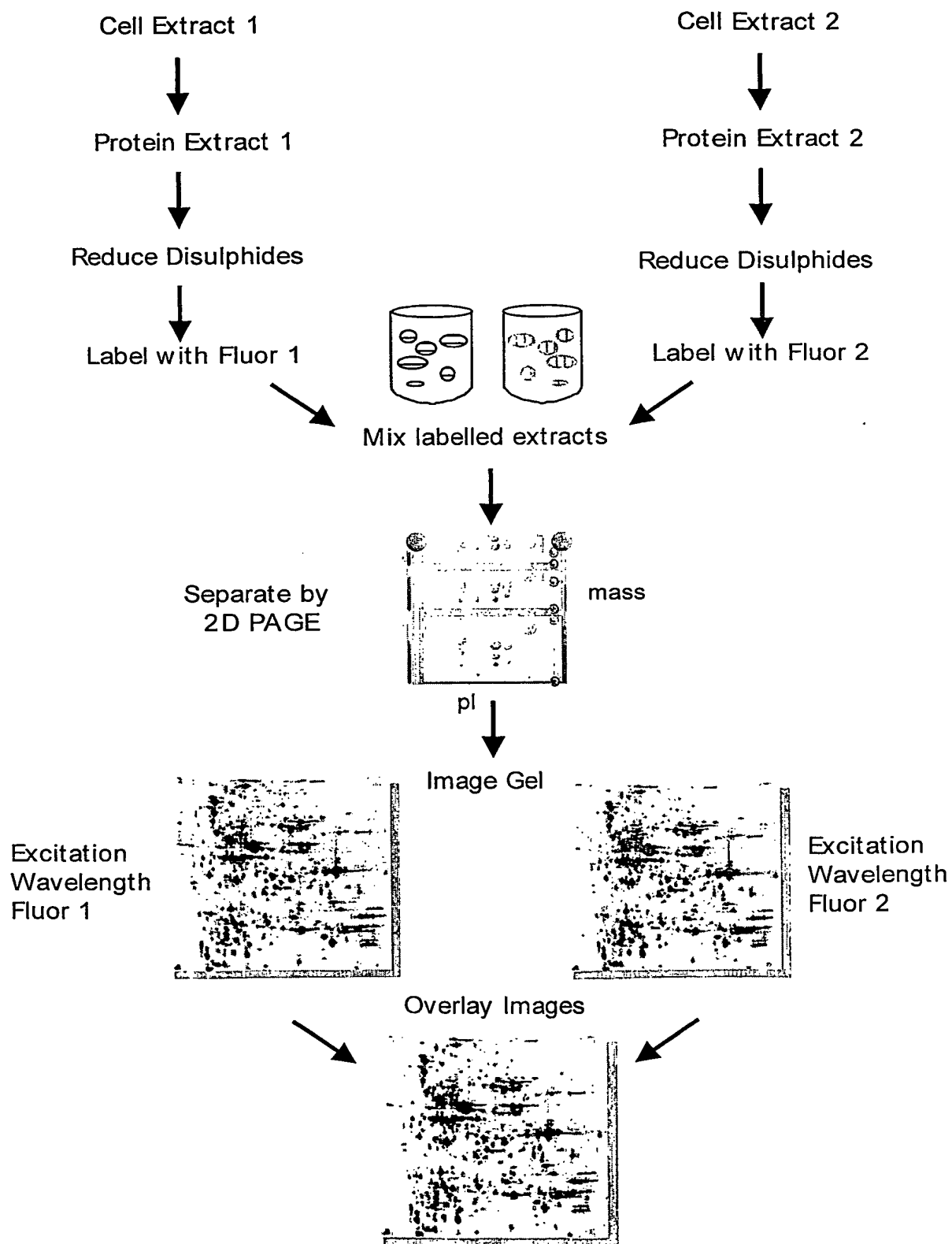
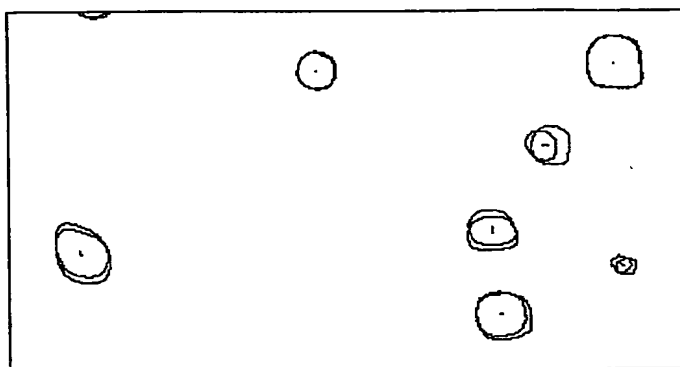
Figure 1/2Schematic Diagram of Example Workflow

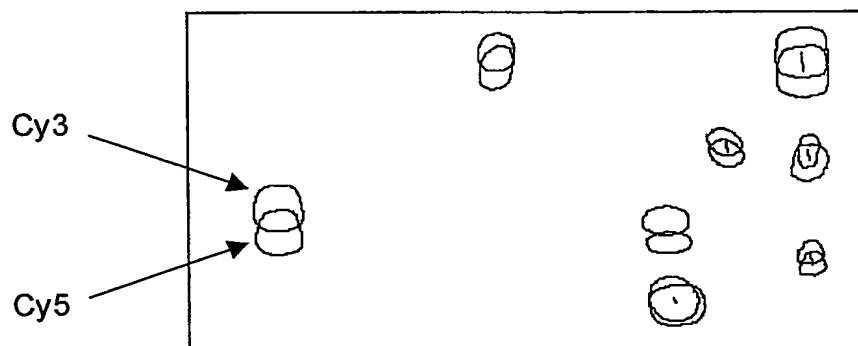
Figure 2/2

Overlay Images of Proteins Labelled with Dye Sets 1, 12 and 9
and Separated by 2D Electrophoresis with Outlines of Labelled
Protein Spots to Demonstrate Positional Matching

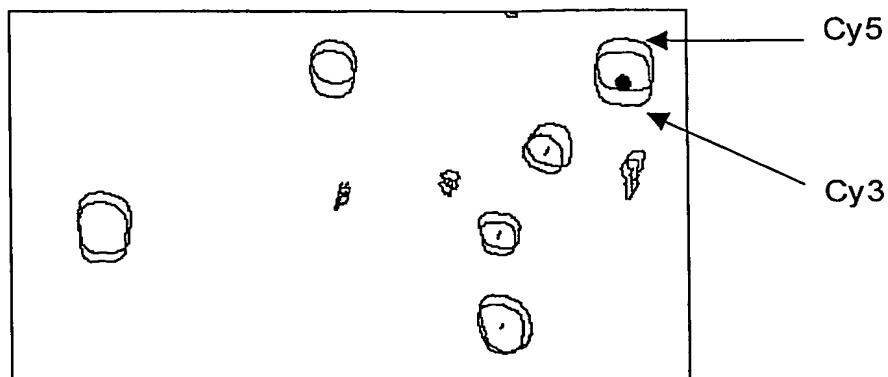
Overlay of Preferred Dye Set 1



Overlay of Dye Set 12



Overlay of Dye Set 9



INTERNATIONAL SEARCH REPORT

International Application No.

PCT/GB 00/0142

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G01N33/68 C07K1/13 C09B23/02 C09B67/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G01N C07K C09B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ERNST L A ET AL: "CYANINE DYE LABELING REAGENTS FOR SULFHYDRYL GROUPS" CYTOMETRY, ALAN LISS, NEW YORK, US, vol. 10, no. 1, 1989, pages 3-10, XP000071370 ISSN: 0196-4763 the whole document	8-14
A	US 5 268 486 A (MUJUMDAR RATNAKAR B ET AL) 7 December 1993 (1993-12-07) column 3, line 18 - line 45; examples 1,2 -/--	1-19



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

17 April 2003

Date of mailing of the international search report

02/05/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Ginoux, C

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 3142

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>US 6 127 134 A (WAGGONER ALAN ET AL) 3 October 2000 (2000-10-03) matched set of dyes. lysine is labelled in a minimal way column 5, line 50 -column 6, line 11 & WO 96 33406 A 24 October 1996 (1996-10-24) cited in the application -----</p>	1-19

INTERNATIONAL SEARCH REPORT

International patent family members

International Application No

PCT/GB 3142

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5268486	A	07-12-1993	US 5486616 A 23-01-1996
			US 5569766 A 29-10-1996
			DE 3912046 A1 15-03-1990
			DE 3943817 C2 12-07-2001
			JP 2191674 A 27-07-1990
			JP 2757965 B2 25-05-1998
			JP 10096727 A 14-04-1998
			JP 2898264 B2 31-05-1999
			JP 10088012 A 07-04-1998
			JP 2002207041 A 26-07-2002
			US 5569587 A 29-10-1996
			US 2002146736 A1 10-10-2002
			US 2002142340 A1 03-10-2002
			US 6225050 B1 01-05-2001
			US 5627027 A 06-05-1997
			US 6048982 A 11-04-2000
US 6127134	A	03-10-2000	AU 709733 B2 02-09-1999
			AU 5557396 A 07-11-1996
			AU 5950099 A 03-02-2000
			CA 2218528 A1 24-10-1996
			EP 0821787 A1 04-02-1998
			JP 11505324 T 18-05-1999
			WO 9633406 A1 24-10-1996
			US 2002177122 A1 28-11-2002
			US 6426190 B1 30-07-2002
			US 6043025 A 28-03-2000